

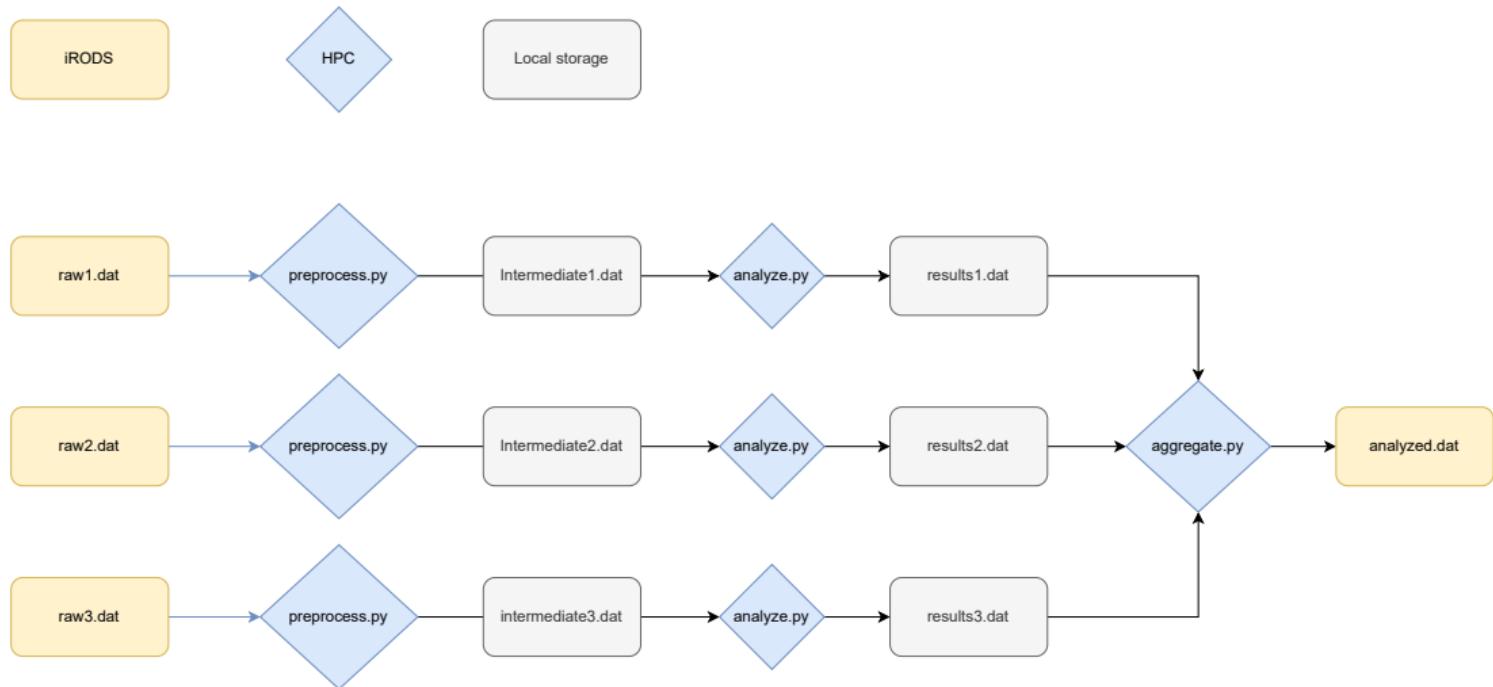
Snakemake

Raoul Schram

November 27, 2025



What is a computational workflow?



What is Snakemake?

- Workflow engine from 2012 (Paper) inspired by Make
- More than 10 citations a week
- Active development
- Open Source, MIT
- Install through conda
- “Syntactical extension of Python”
- Steps through input and output files
- Support for different execution backends (SLURM) and remote storage (S3)
- Dynamic workflows (checkpoints)

```
rule myrule:  
    input:  
        "path/to/inputfile",  
        "path/to/other/inputfile",  
    output:  
        "path/to/outputfile",  
        "path/to/another/outputfile",  
    shell:  
        "somecommand {input} {output}"
```

Presents from Sinterklaas (exercise)

- Sinterklaas asks his Pieten to distribute 10 million packages
- Problem: the Pieten have an authority problem
- Atom Piet: only indivisible numbers
- Professor Piet: only numbers that end with three
- Average Piet: only numbers that have an average decimal of 5
- Drunk Piet: only numbers that have repeated numbers (e.g. 11 or 22)
- Astrology Piet: No numbers that end with nine
- Lazy Piet: Everything else



How big is the pile of packages in the canal?



iRODS, Yoda and iBridges

- iRODS is a datamanagement platform, at UU we have Yoda
- Not only files, but dataobjects: data + metadata
- Many ways to access it: iCommands, Yoda portal, GoCommands, iBridges

The screenshot shows the Yoda Portal homepage. At the top left is the Utrecht University logo. On the right is a yellow "Sign in" button. Below the header is a black bar with the text "Yoda Portal". The main content area features a photograph of four people (three women and one man) working together around a table, looking at a map or document. To the right of the photo is the Yoda logo (a stylized yellow sun-like icon) and the text "Yoda research data management". Below that is the text "Welcome to Yoda!" and "Yoda is a share-collaborate environment for research data." A yellow "Help" button is located in the bottom left corner of the content area. In the bottom right corner of the content area is the Utrecht University logo. At the very bottom of the page are navigation icons.

Snakemake iBridges plugin

- <https://github.com/qubixes/snakefile-storage-plugin-irods>
- Branch: switch-to-ibridges
- New feature: wildcards
 - ▶ Can target all files with a certain pattern
 - ▶ `/zone/user/analysis/{run_id}/temp.txt`
- New feature: metadata support
 - ▶ (iRODS) metadata is not exposed through the snakemake interface
 - ▶ `/zone/user/analysis/data.txt.metadata.json` or
 - ▶ `/zone/user/analysis.metadata.json`
 - ▶ Read or write .json files to interact with metadata
- New feature: integrate with iBridges CLI
 - ▶ Re-use cached passwords from multiple iRODS environments
- New Feature: relative paths

Exercise part 2

- Modify the previous excercise to use an iRODS server
- pip install
git+https://github.com/qubixes/snakefile-storage-plugin-irods.git@switch-to-ibridges
- Username/password: test[i]
- irods_environment.json:

```
{"irods_home": "/tempZone/home/test1",
"irods_host": "irodsserver.researchcloudde.surf-hosted.nl",
"irods_port": 1247,
"irods_user_name": "test1",
"irods_default_resource": "trainingResc",
"irods_zone_name": "tempZone" }
```

```
1 storage:
2   provider="irods",
3   environment_file="/home/qubix/.irods/irods_env_surf.json",
4   cwd="/uu/home/research-christine",
5
6
7 rule all:
8   input:
9     |   meta=storage("irods://~/research-christine/test_plugin/test.txt.metadata.json")
10
```

